DOCUMENT RESUME

ED 086 986 CS 200 921

AUTHOR Lesgold, Alan M.
TITLE Variability in Children's Comprehension of Syntactic Structures.
INSTITUTION Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 73
NOTE 16p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Comprehension; *Elementary Grades; Language Research; Linguistics; *Pronouns; *Reading Comprehension; Reading Research; *Sentence Structure; *Syntax

ABSTRACT

In this report data are presented that challenge the difficulty ordering for anaphoric syntax (e. g., pronouns) proposed by Bormuth, Manning, Carr, and Pearson in 1970. It is suggested that any such difficulty ordering resulting from tests of the form proposed by Bormuth (1970) will have uncontrolled variability due to semantic factors that have yet to be carefully analyzed and controlled. (Author)

LEARNING RESEARCH AND DEVELOPMENT CENTER

1973/10

VARIABILITY IN CHILDREN'S COMPREHENSION OF SYNTACTIC STRUCTURES

ALAN M. LESGOLD

VARIABILITY IN CHILDREN'S COMPREHENSION OF SYNTACTIC STRUCTURES

Alan M. Lesgold

Learning Research and Development Center
University of Pittsburgh

1972

The research reported herein was supported by the Learning Research and Development Center, supported in part by funds from the National Institute of Education (NIE), United States Department of Health, Education, and Welfare. The opinions expressed do not necessarily reflect the position or policy of NIE and no official endorsement should be inferred.

Abstract

Data are presented that challenge the difficulty ordering for anaphoric syntax (e.g., pronouns) proposed by Bormuth, Manning, Carr, and Pearson (1970). It is suggested that any such difficulty ordering resulting from tests of the form proposed by Bormuth (1970) will have uncontrolled variability due to semantic factors that have yet to be carefully analyzed and controlled.

[illegible]

[illegible]
[illegible]

Bormuth, Manning, Carr, and Pearson (1970) have reported results suggesting that stable difficulty orderings could be obtained for three classes of syntactic forms containing a total of 55 separate variants. There were difficulty differences both between and within the three classes (intra-sentence, inter-sentence, and anaphora[2]), from which they argued that these categories may be related to stages in a learning hierarchy for syntax comprehension skills. Further, there have been recent suggestions that additional experimentation of this sort, combined with linguistic analysis, will result in discovery of such a hierarchy (e.g., Carroll, 1972; Fraser, 1972). The present results, a replication of part of the Bormuth et al. work, suggest that this approach is unlikely to succeed unless such factors as semantics and constraints on information processing capacity are concurrently considered.

[illegible]

---

[1] The experiment was conducted by Hildrone DeGood. Karen Block and Charles Perfetti read an earlier draft and helped to improve it.

[2] Anaphora is the term used to denote a structure in a sentence, e.g., a pronoun, that derives its meaning from a previous sentence or an earlier part of the present sentence.

written and then a second was constructed with some reference back to the first (e.g., John went to the store. He bought a pear.). A paragraph was then constructed around the two-sentence cluster. A question was generated for the paragraph by substituting a wh-word (Bormuth, 1970) for the anaphora (e.g., Who bought a pear?). Finally (for anaphora only), multiple-choice alternative answers were written. The item thus consisted of a paragraph, a question, and alternative answers to choose from.

There is a basic problem in using procedures of this sort for measuring the difficulty of one syntactic structure relative to another. This is the confounding of syntax with semantics. Consider a potential comparison between a personal pronoun structure (e.g., he) and a pro-clause form (in which that or so might stand for an entire clause). The two syntactic forms occur in different semantic contexts: There is no syntactic transformation series that will produce a personal pronoun from the semantic structure underlying a pro-clause. Now this does not necessarily mean that a difficulty ordering such as that of Bormuth et al. should be discounted. After all, the confounding of syntax and semantics may, for such purposes, be complete.

If the Bormuth et al. data are consistently replicable, then we can still look for a learning hierarchy related to that data, even though we would not know, initially, the relative roles of syntax and semantics. On the other hand, it may be that the difficulty ordering is not stable, that the particular confoundings of syntax and semantics in the item forms of Bormuth et al. are only partial. Then, if semantic content is not controlled, spurious measurements of syntax processing ability may result. If the latter is the case, then a difficulty ordering is premature until a semantic analysis as complete as Bormuth's (1970)

syntactic analysis is available. The present results are a demonstration that the difficulty ordering of syntactic structures is not stable, implying that such further analyses are necessary.

There is another problem in measuring syntax difficulty. Sometimes the answer to a question may be betrayed by semantic constraints. The paragraph may contain only one semantically possible answer to the question. For example, a who question after a paragraph with only one animate noun could be answered without the use of any skills related to the target syntactic structure. Such an extreme problem is not likely in a careful study such as that of Bormuth et al. More generally, though, semantic constraints on the answer to an item's question that result from choice of content words for the item's paragraph are difficult to determine and may not be completely controlled in studies such as that of Bormuth et al. and the present experiment. Again, this may be a moot point if such constraints are perfectly correlated with syntax, but the present results rule out this possibility.

In addition to semantic difficulty and the extent to which the paragraph "gives away" the answer, there is a third potential source of variance in the difficulty ordering produced by the Bormuth et al. method. Two passages that have a target syntactic structure in common may differ in the extent to which they can otherwise be processed to the point at which the critical structure is relevant. One may sensibly hypothesize, for example, that an anaphora cannot be comprehended unless both it and its antecedent are simultaneously in operating (short-term) memory. For example, Bormuth et al. found that items like (1) are comprehended 87 percent of the time, while items like (2) are comprehended only 66 percent of the time, by fourth-graders. If he had used a different pro-clause item like (3), instead of (1), perhaps only half the children

would have understood pro-clause forms, thus reversing the ranking of clause demonstratives and pro-clauses (so). Passage 3 is longer than (1) and may be more difficult to encode due to its idiomatic content. Most important, so stands in place of a much more complex construction in (3) than in (1). Unless a nine-year-old child is able to record splice the mainbrace into a single image, he may lack the ability to process that phrase to the point of being able to resolve the anaphoric reference. No control procedure is available to insure that Bormuth's ranking arises only from syntax differences and not from differences in passage wording or the amount of processing required to get from surface to underlying structure.

(1) Joe may cry. If so, the rest of us will be sad.

(2) Joe is coming home. That is the best news I've heard all week.

(3) Joe may splice the mainbrace. If so, the rest of us will be glad.

The present results arise from what was expected to be a screening task for an experiment on memory for syntax. Thus, they provide information about only a subset of nine out of Bormuth et al.'s fourteen anaphora forms. The differences between the two studies are: (1) the present study used oral, constructed responses while Bormuth et al. used written, multiple-choices responses; (2) the present study explicitly controlled the number of semantically plausible potential answers in each passage; and (3) the location of the target structure in the passage was counterbalanced in the present study.

Method

Subjects. Forty students from a campus laboratory school and 40 students from an urban public school participated as subjects. The campus group ranged in age from 8.0 to 10.0 years with a median of 8.8, while the urban group ranged from 8.0 to 11.5 years with a median of 9.4. All were in third or fourth grade.[3]

Materials. Fourteen anaphora forms were tested in this study. They are listed with examples, in Table 1. Three items were written for each form. An item consisted of a paragraph plus a question. The paragraph, in turn, consisted of two filler sentences plus a two-sentence critical structure sequence. The first of the two critical sentences contained an antecedent which was referenced by an anaphora in the second. Each of the three items for a form had a different location for the critical sentences in the paragraph: before, between, or after the filler sentences. Each item's passage was constructed so that there were two semantically sensible answers to the question. The correct choice was determined by the anaphora syntax. The question for each item was written by substituting the appropriate wh-word for the anaphora and then applying the shortest sequence of transformations that would turn the anaphora sentence into a question.

Each question was typed on a 21.6 x 27.9 cm sheet in 0.42 cm gothic type and covered with a clear plastic sleeve. The 42 pages were presented to subjects in a loose-leaf binder. Order of occurrence of the items was approximately counterbalanced.

---

[3]The campus laboratory school actually had two-year groupings, one of which was equivalent to third and fourth grades.

Table 1

The Relative Difficulty of Anaphoric Structures

Anaphora Type	Example	Percentage Correct		
		Bormuth et al.	Present Study	Difference
Personal Pronoun	Joe left the room. He had....	64.5	91.7	+27.2
Relative Pronoun	The man who lives next door....	82.8	78.3	− 4.5
Noun-Phrase Demonstrative	The old dog belongs to Joe. That is his....	81.5	75.0	− 6.5
Negated Pronoun	The gang went shopping. No one bought....	67.4	67.5	+ 0.1
Pro-Verb, so-do	John likes Mary. So does Bill.	82.8	60.8	−22.0
Pro-Verb, so-be/have	Joe is sick. So is Bill.	76.1	54.6	−21.5
Pro-Adverb Locative	I am upstairs. It is cold there.	---	79.2	---
Pro-Clause, so	Joe may go. If so, we will....	86.8	64.2	−22.6
Clause Demonstrative	Joe is dead. That leaves two of us.	66.3	90.0	+23.7
Pro-Adjective	John is careful. Bill is that way, too.	---	66.2	---
Semantic Substitute–Noun	Those dishes are expensive, but this china....	65.5	82.9	+17.4
Semantic Substitute–Verb	Jim shot first. John fired, too.	---	82.5	---
Semantic Substitute–Clause	Bill went to the bank. That trip made us short one outfielder.	---	75.0	---
Semantic Substitute–Adverb	John works carefully. Bill also works precisely.	---	70.8	---

Procedure. The procedure was first explained to S. Each S worked individually and at his own pace reading each paragraph and then orally answering the question. The Ss from each school were split into two groups of 20 each. One group read the passages (generally silently), while the other group followed along reading (silently) while they heard a tape recording of the passage as well. After completing 30 of the 42 passages, each S stopped for a brief cookie break. Occasionally, S would respond to a question by stating the anaphora word (e.g., he). The subject was then asked to be more specific (e.g., Whom do you mean by he?).

Scoring. Each response to a passage by a given S was punched onto a separate computer card. A list was then produced, sorted by item. This made it possible to examine all answers to an item at once. Each answer was scored as correct only if the verbatim antecedent of the anaphora in the given passage was stated or implied by a stated synonym of superordinate category term. Answers referring to more entities besides the correct antecedent were also counted wrong. Scoring was verified by a second observer.

Results

Two reliability measures were computed. The cell means for the fourteen anaphora types were compared for the two schools (r = 0.94, df = 12, $p < .001$) and for the two presentation conditions (r = 0.88, df = 12, $p < .001$). Both correlations are quite high, suggesting that these results are relatively stable for the particular items used.

To determine whether the anaphora forms differ in their difficulty, a 20 (Subjects) X 14 (Anaphora) X 2 (Schools) X 2 (Presentation mode) partially nested analysis of variance was performed on the scores (zero to three correct) of Ss on each of the anaphora forms. There were significant differences among the anaphora forms ($F = 19.6$, $df = 13/988$, $p < .0001$) and no interactions of Anaphora with Schools or Presentation mode (F's ≤ 1.28, $df = 13/988$). The percentage correct for each of the anaphora types is shown in Table 1.

Schools was a significant variable ($F = 27.5$, $df = 1/76$, $p < .0001$), with the campus lab school ($M = 2.42$) showing higher performance than the urban public school ($M = 2.04$). There is alsc some reason to suspect that while the urban school Ss performed equally under the two presentation conditions ($Ms = 2.03, 2.04$), the campus lab school Ss benefited from the read-and-hear condition ($Ms = 2.54, 2.29$; $F = 2.89$, $df = 2/76$, $p < .10$).

The important result is the comparison of the rank orders of difficulty for those nine anaphora conditions common to both the present study and that of Bormuth et al. (1970). The mean percent correct for each anaphora form in each study can be seen in Table 1. There is a significant negative correlation between the two sets of means (Spearman rho = 0.66, $N = 9$, $p < .05$).

Discussion

Since Bormuth et al. (1970) suggest that their difficulty ordering may be the basis for a learning hierarchy for anaphora syntax, it is important to be certain that their data and the present data cannot possibly

be a basis for the same hierarchy. The negative rank correlation gives some evidence in this regard. However, it would be possible to find such a correlation if one subject sample possessed the skills needed to comprehend Structure A, while a second subject sample had not yet acquired those skills. If both samples performed at a middle level on Structure B, then the first would display a rank ordering of A > B, while the second had an ordering B > A. Since many structures have a large set of interpretation rules, the relative order of difficulty could change as rules with differing importance are acquired.

This is not the case for the present data since some forms, such as the two pro-verb forms and the pro-clause (so) form, are better comprehended by the Bormuth et al. Ss, while others, such as personal pronoun, are better comprehended by the present Ss. This complete crossover requires us to reject any argument that the differences between results of the Bormuth procedure and the present study are due to developmental differences in the two studies' subject pools.

The present study was intended as a screening task for an experiment derived from Bormuth's work. Thus, every effort was made to reproduce any potentially critical procedures of Bormuth's. The results obtained are statistically reliable. We cannot help but conclude that syntax is not the basis of a hierarchy of comprehension skills that have not been completely acquired by the time children are in fourth grade. Any structure for which high performance is shown in either the present study or that of Bormuth et al. must be assumed not to present a straightforward syntactic hurdle to fourth-grade children.

Why, then, do children perform poorly on a particular structure on even one of the two tests. There are two potential reasons: (1) the child may not know the interpretation rules required to understand the

structure in a particular semantic context; or (2) he may lack the real-time capacity for applying those rules. There is a good reason to believe that both of these factors play a role but that the second is of predominant importance for children as old as the present Ss. The available evidence is primarily restricted to studies of pronominalization.

Fredrick, Golub, and Johnson (1970) have found that multiple-choice pronoun items similar to those of Bormuth et al. are correctly answered only 28 to 50 percent of the time. The items are grammatically similar to the personal and relative pronoun items of the present study, which are correctly answered 92 and 78 percent of the time, respectively. The apparent difference is in the level of semantic ambiguity in the two cases. Consider the following examples, from Fredrick et al. and the present study, respectively:

(1) The notebook on her desk covered up my drawing which was very messy. [Identify which.]

a. notebook

b. desk

c. covered up

d. drawing

e. messy

(2) Two men were walking down the street. One man had on a hat. The big man who was standing on the corner is my father. Who was standing on the corner?

The first example is semantically more complicated. It requires that the child realize that the "closest semantically acceptable antecedent" rule applies. There are three semantically acceptable potential

antecedents, while there are only two potential answers to (2). Even children who get (1) and (2) correct will likely waver in handling (3).

(3) I put the package on the table. Because it was tilted, it fell off. [Identify it.]

a. package

b. table

c. tilted

d. fell

e. I

The point is that grammatical rules for nine-year-old children are not abstract structures that apply mechanically. They are inextricably bound up with semantics (cf. Palermo & Molfese, 1972). The presence of a subset of the pronoun interpretation rules in even five-year-olds (Chomsky, 1969) does not mean that these rules operate abstractly and free of semantic influence. The potential success of syntax tests with uncontrolled semantics is low in general; it is practically non-existent for pre-formal-operations children.

The second potential reason for variability in tests of syntactic competence is differences in the extent to which a given item exceeds or stays within the child's channel capacity as an information processor. Recent findings by the present author[4] suggest that adult-like comprehension of personal pronoun sentences is more likely when a child can replace cumbersome surface-structure segments with imaginal codes. Thus, imagery factors apparently play a role in comprehension performance, also.

---

[4]Lesgold, A. M. Memory structure for pronoun sentences: Imagery effects on comprehension channel capacity, in preparation.

These arguments suggest that the design of a syntax comprehension curriculum cannot readily be based upon tests of the sort described here and by Bormuth et al. Such tests have uncontrolled variance due to imagery and semantic factors. The structures tested are often (but not always, cf. Palermo & Molfese, 1972) understood by children when presented in sufficiently semantically simple circumstances. On the other hand, sufficiently complex semantics will probably override knowledge of any structure in young children. Perhaps semantic analysis, at the level of care that Bormuth (1970) has brought to syntactic analysis, will be useful in the design of instruction. However, at the present, the use of difficulty orderings for syntax without regard to semantics is not likely to lead to improved instruction in comprehension.

It is certainly possible that abstract syntax rules are being acquired during the ages from which Bormuth's and the present samples were drawn. Further, it is conceivable that one might want to test for the presence of such rules. This could be done by using nonsense words instead of lexical words in the various item forms used in this study. This would yield measures of "pure" syntax ability. However, many syntactic rules interact with semantics, so this "pure" ability might not be very relevant.

References

Bormuth, J. R. On the theory of achievement test items. Chicago: University of Chicago Press, 1970.

Bormuth, J. R., Manning, J., Carr, J., & Pearson, D. Children's comprehension of between- and within-sentence syntactic structures. Journal of Educational Psychology, 1970, 61, 349-357.

Carroll, J. B. Defining language comprehension: Some speculations. In J. B. Carroll and R. O. Freedle (Eds.), Language comprehension and the acquisition of knowledge. Washington, D.C.: V. H. Winston & Sons, 1972.

Chomsky, C. The acquisition of syntax in children from 5 to 10. Cambridge: MIT Press, 1969.

Frase, L. T. Maintenance and control in the acquisition of knowledge from written materials. In J. B. Carroll and R. O. Freedle (Eds.), Language comprehension and the acquisition of knowledge. Washington, D.C.: V. H. Winston & Sons, 1972.

Fredrick, W. C., Golub, L. S., & Johnson, S. L. Analysis of the Linguistic Ability Test, grades 4 & 6. Madison, Wisconsin: Wisconsin Research and Development Center for Cognitive Learning, 1970. (Technical Report No. 121)

Palermo, D. S., & Molfese, D. L. Language acquisition from age five onward. Psychological Bulletin, 1972, 78, 409-428.